High Wind and Energy Specific Models for Global Production Forecast

Carlos Alaíz, Álvaro Barbero, Ángela Fernández, José R. Dorronsoro *

Dpto. de Ingeniería Informática and Instituto de Ingeniería del Conocimiento

C. Tomás y Valiente 11, Universidad Autónoma de Madrid,

28049 Madrid, Spain

Abstract

High and low production regimes are in principle different enough as to warrant the use of regime–specific models for the prediction of wind energy production. The simplest way to identify them may be the use of concrete wind/production thresholds. The computation of these thresholds requires an estimate of the most likely future regime and in this work we consider both NWP wind speed forecasts and also the production forecasts of a global full operation range model. As we shall illustrate over the aggregated wind energy production of a very large area of Spain, a production threshold–based approach gives consistently better results than the alternative, wind speed based, procedure.

1 Introduction

Wind energy usually spans two opposite production regimes of high and low energy generation. Given the wind speed distribution, it can be said that the low energy regime represents the most standard situation. It is certainly of great interest for individual farm producers, as it may be used to determine maintenance periods and also because the corresponding relative errors, that may have a big influence in market deviation penalties, are usually largest at the low energy regime. On the other hand, and from the point of view of transmission and system operators (TSOs) that take care of large wind producing areas, high energy regimes may be still rarer than for individual farms. However, the ever increasing wind installed power and market penetration imply that large

*Corresponding author; email: jose.r.dorronsoro@iic.uam.es; phone: +34 91 497 2329, fax: +34 91 497 2334.

production variations may have a large impact on the levels of spinning reserve to be maintained and on the concrete measures to be taken to ensure the overall stability of the electrical grid.

This makes very important the accurate wind energy prediction of high production situations. However, this may be hard for global models built across the full energy production range that TSOs typically use. In fact, the Weibull distribution of wind speed implies that it will be dominated by low and medium values and, therefore, the same will happen for wind energy. Thus, the models typically used for global energy prediction, such as neural networks, will be dominated by lower wind speed samples and, therefore, high wind energy (HWE) forecasting may be less accurate, as it will be under–represented in training samples. This makes quite interesting the construction of specific models for HWE forecasting.

In this work we will address this issue in the concrete setting of Spain's global wind energy forecasting. Observe that before applying a given high or low energy model, one has to detect which regime will most likely be true, something that necessarily requires a regime forecasting methodology. Two different approaches open themselves naturally. The first one is to rely on NWP wind forecasts and to determine that a HWE regime will hold whenever a certain high wind speed threshold is surpassed. The second alternative would be to define HWE situations in terms of the production levels themselves. However, while high wind situations could be directly derived from numerical weather predictions (NWP) available with ample anticipation, future high production situations cannot be directly derived from third party NWP values and ad–hoc models will have to be built.

Good wind speed or wind production forecasts are crucial for each one of these approaches. At first sight one might think that the use of high wind predictions would be safer, for if NWP forecasts are wrong, it is very likely that so will be energy production forecasts derived from them. In fact, this could be a sensible observation when dealing with individual farms. However, in a wide area situation, one has first to choose a single wind speed value that is representative of the whole area, something that may get harder the bigger the area under consideration. On the other hand, this is precisely what a global production model does: to integrate concrete NWP forecasts for the several NWP grid nodes considered into a single production value. Thus, such a model may better capture the global influence of the many individual NWP nodes on the overall energy production.

As we shall numerically illustrate, the second approach will be indeed more successful, providing better forecasts for both the overall full range productions and the HWE situations while, on the other hand, a high wind approach, at least as done according to the methods to be presented later, does not seem to offer advantages over a single full range model.

The paper is organized as follows. In sections 2 and 3 we shall elaborate on the above alternative characterizations of HWE situations, discussing also how to determine optimal wind speed or production thresholds. We shall apply both approaches in section 4 to the prediction of the overall wind energy production of Spain. As we shall see, it may be possible to build in a semiautomatic fashion high production–specific models with a better performance than that of full production range models. On the other hand, high wind–specific models may outperform full range ones, but only after a careful selection of threshold values; in particular, and as it results from our experiments, automatic approaches to that goal may result in worse models. The paper ends with a brief discussion and conclusions section.

2 Wind Speed–based HWE Models

Intuitively there is a clear relationship between high wind situations and high energy production. Thus, the simplest approach to build specific predictors is to define a wind speed threshold value and to construct different energy prediction models for NWP forecasts whose absolute speed falls above and below it.

In a global production situation, a first difficulty is to combine the wind speeds at the various nodes of the NWP grid covering the geographical area for which global production forecasts are wanted. In principle it would be desirable to analyze first local production data and, then, to integrate them on a global setting. However, this would require detailed information of local, de–aggregated wind energy production, something that may not be always possible. A simpler way to proceed is to work with a global weighted wind speed, where the wind speeds at the NWP grid nodes close to each wind farm are weighted by the farm's rating. In other words, if we have M farms, v_i is the wind speed forecast for the node closest to wind farm i, $1 \leq i \leq M$, and P_i is its installed power, we compute the weighted wind speed as

$$\overline{v} = \frac{1}{\mathcal{P}} \sum_{1}^{M} v_i P_i,$$

where $\mathcal{P} = \sum_1^M P_i$ is the total installed power of the area under consideration. Figure 1 gives the histogram of the distribution of these $\overline{v}$ values. As it can be seen, it has a shape similar to that of a Weibull distribution, which is usually taken as representative of actual site wind distributions. We can compare the weighted wind speeds to the corresponding productions. This is done in figure 2; while somewhat reminiscent of the typical load curve of an aerogenerator, it also has clear differences. For instance, it is much more concentrated at low and medium weighted wind speeds, while spreads are clearly higher for high wind speeds.

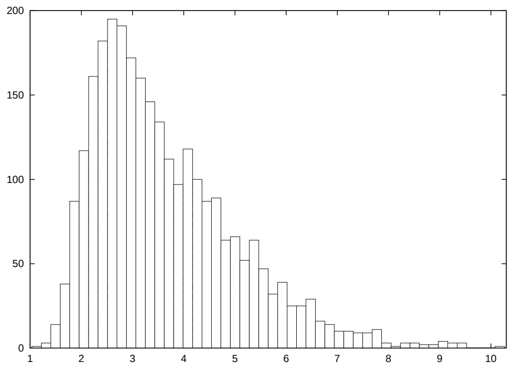

Figure 1: Histogram of the weighted wind speed forecasts.

This lends support to the use of different models for low and medium speeds on the one hand, and high speeds on the other. The crucial question is to decide on the appropriate speed threshold $\overline{v}_0$; another aspect to discuss would be how to implement the transition between both models. For simplicity we will completely separate them, using one model for NWP forecasts with a weighted wind speed $\overline{v}$ such that $\overline{v} < \overline{v}_0$ and another one when $\overline{v} \geq \overline{v}_0$. A possibly more sensible option could be to consider a transition zone $(\overline{v}_0 - \delta, \overline{v}_0 + \delta)$, use one model when $\overline{v} < \overline{v}_0 - \delta$, another one when $\overline{v} > \overline{v}_0 + \delta$ and a combination of both when $\overline{v}_0 - \delta \leq \overline{v} \leq \overline{v}_0 + \delta$. However this would require us to compute both parameters δ and $\overline{v}_0$ and, thus, would be costlier if done in an automatic fashion (and quite a bit messy if done on a heuristic trial–and–error basis). Hence, we shall only consider the first option.

Therefore, assuming a threshold value $\overline{v}_0$ has been fixed, we first split the training sample $\mathcal{S}$ into two subsets $\mathcal{S}_L = \{X_i : v_i < \overline{v}_0\}$ and $\mathcal{S}_H = \{X_i : v_i \geq \overline{v}_0\}$, where v_i denotes the weighted wind speed associated to the NWP pattern X_i and, then, build a low speed model F_L over $\mathcal{S}_L$ and a high speed one F_H over $\mathcal{S}_H$. When we want to forecast the energy production from a new NWP pattern X_j, we simply apply either F_H or F_L depending on whether v_j is above or below $\overline{v}_0$.

We consider next how to determine the optimal $\overline{v}_0$. There is a general machine learning methodology for obtaining a possibly optimal parameter values which consists to split the data available for modeling into a training subset $\mathcal{T}$ and a validation one $\mathcal{V}$, to construct several models over $\mathcal{T}$ under changing values of $\overline{v}_0$ and to settle into the one that gives a smallest error over $\mathcal{V}$. The trickiest part is how to generate the successive threshods to be tested. There are many proposals for this, many of them centered on evolutionary procedures that stochastically explore the available range for $\overline{v}_0$ using the prediction error over $\mathcal{V}$ as the fitness function. One of the best known among these is the (μ, λ) Covariance Matrix Adaptation–Evolution Strategy (CMA–ES), proposed by Hansen and Ostermeier [2]. CMA–ES is a very effective black–box optimizer, particularly when several

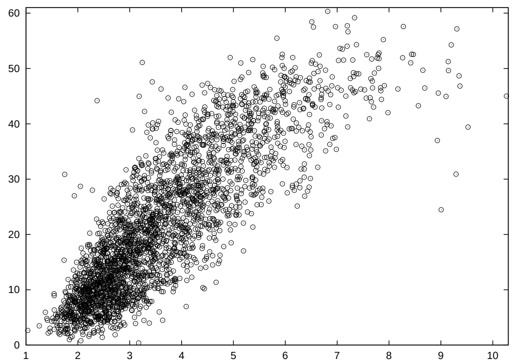

Figure 2: Weighted wind–production relationship.

parameter values have to be obtained simultaneously and they may range over a large region. In the present situation, however, a single wind speed value has to be obtained and range limits are easily defined. Thus, we shall simply apply a one–dimensional search over a discretization of the possible wind speed values, bounding the search region below by the average weighted wind speed. No upper bound will be considered. This has the risk that the high wind speed subsample $\mathcal{S}_H$ be empty (as it will be the case in some of our experiments), which implies that only a single global model will be built and applied to new NWP forecasts. This issue will be considered in further work.

3 High Energy Production–based HWE Models

As said before, while high wind speed is clearly an important contribution to high wind energy for localized regions, the relationship between a single wind speed value and the wind energy production of a large area may be more involved. On the other hand, it is easy to identify a high energy regime by applying properly selected production thresholds. Figure 3 gives the production histogram for the selected period, which is rather different from the wind speed histogram. Therefore, we shall consider here whether it is possible to somehow classify NWP predictions as being likely or not to result in a high wind energy production and, if so, whether one can take advantage of this classification to derive better production forecasts.

There are several approaches we could follow. Since we are, in fact, solving a classification problem, one possibility is to label the NWP pattern X_i of a given pair (X_i, p_i) with p_i the actual production value corresponding to X_i as either high– or low–production according to whether p_i lies above or below a certain production threshold p_0, construct then a classifier that assigns new NWP forecasts X to one of these two classes and apply then on X an appropriate high or low

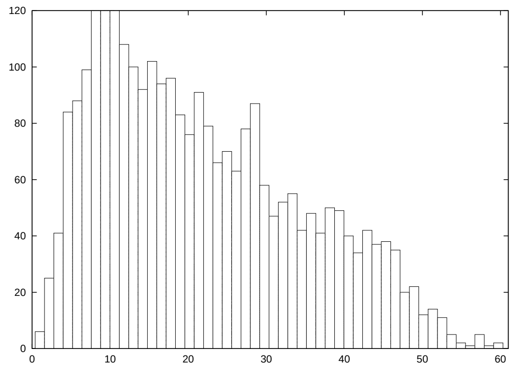

Figure 3: Histogram of wind energy production

production model.

While sensible, this approach has the drawback of mixing classification and prediction models with a different nature. A more straightforward way to proceed is to simply use a global production model $F(X)$ to identify a HWE situation depending on whether we have have $F(X) \geq p_0$ for a given NWP pattern X. This means that at a given moment we must have three prediction models, a global one $F_G(X)$ and two specialized ones, a low energy model $F_L(X)$ and a high energy model $F_H(X)$. Assuming a threshold value p_0 has been fixed, one first builds F_G over the full training sample $\mathcal{S}$ and then splits its patterns into two subsamples, $\mathcal{S}_L = \{X_i : F_G(X_i) < p_0\}$ and $\mathcal{S}_H = \{X_i : F_G(X_i) \geq p_0\}$. The low energy model F_L is built next over $\mathcal{S}_L$ and the high energy one F_H is built over $\mathcal{S}_H$. Similarly, when we want to forecast the energy production from a new NWP pattern X_i, we simply apply first F_G and then either F_H or F_L depending on whether $F_G(X_i)$ is above or below p_0.

Notice that here the two models that are applied to a given X_i are similar in nature; moreover, if the prediction errors of the global F_G are low, we can expect the output of the global model to give a reasonable indication of the high energy regime to be expected. Moreover, we may also expect the global model F_G to establish a connection between NWP values and actual predictions which will be stronger than the one possible using wind speed forecasts. As we shall see in the next section, this is indeed the case, and this approach will yield better results.

Finally, since we also have to determine here a single optimal threshold value, we shall simply apply a one–dimensional search over a discretization of the possible wind speed values. As before, we bound the search region from below by the average production over the training set and will not consider an upper bound. Again this has the risk that the high wind speed subsample $\mathcal{S}_H$ be empty, although this will nor happen in our experiments.

4 Numerical Experiments

In this section we will separately compare the performance of the two previous approaches to high wind energy predictions against a single MLP model to provide wind energy forecasts across all NWP or energy production situations. We shall apply these models to the forecasting of the aggregated wind generation of a large number of wind farms in Spain with an installed power of about 11.500 MW (Spain's total is, as of this writing, more than 16.500 MW).

The forecasting target are the hourly productions for day $D+1$ that must be given before 10:00 AM of day D, the bidding hour of Spain's energy markets. The predictive variables to be used are the surface values for the U and V wind components, the absolute wind speed v, and the pressure P and temperature T forecasts provided by the ECMWF at the 0.50^o resolution at 0000 UTC. About 90 ECMWF nodes are needed to cover the farm area and thus, input dimension is in principle about 450. This has to be reduced, for which principal components are applied to the NWP data retaining 99% of total variance.

We shall work with historic data from July 2005 to June 2007 and use as a testing period the twelve months from July 2006 to June 2007. The models to be used will be built using one year of historic NWP and production data. For each day, four NWP forecasts at hours 00, 06, 12 and 18 will be considered. This results in training sets with 1,460 patterns.

As explained above, our approach requires to determine the high wind and high energy threshold parameters. In both cases this will be done using a discrete search on the allowable parameter ranges, that comprise all values above the averages of weighted wind speed and energy production. To test the effectiveness of a possible threshold, its associated models will be evaluated by their prediction errors over a five–fold validation subset. This means that the historic NWP and production data set will be randomly split into 5 subsets and, in a rotating fashion, four of them will be used as training sets and the remaining one as a validation subset. The fitness function will then be the average over the just described 5 validation subsets of the absolute prediction errors of the combination of the low and high speed/energy models.

Although the optimal parameter search is implemented using a focusing procedure for more efficiency, it would still be too time consuming to be done on a daily basis. In our experiments optimal thresholds will be computed once a month over data up to the last day of the previous month and will be kept fixed throughout the new one. Principal component dimensionality reduction will also be done once a month and the projection matrix so obtained will be kept fixed until the next month. On the other hand, all models will be daily updated, so that they keep track of underlying potential production changes. The number of units on the single hidden layer will be 5 in all cases. While

Month	FR MLP	SWR MLP		SPR MLP	
	Full	Split	Improv. %	Split	Improv. %
July	3.35	3.35	0.00	3.29	1.79
August	3.91	4.27	-9.21	3.74	4.35
September	4.14	4.11	0.72	3.88	6.28
October	4.26	4.41	-3.52	4.07	4.46
November	4.23	4.12	2.60	4.19	0.95
December	4.89	4.89	0.00	4.38	10.43
January	4.44	4.87	-9.68	4.04	9.01
February	4.69	4.68	0.21	4.50	4.05
March	4.76	5.32	-11.76	4.28	10.08
April	3.73	3.74	-0.27	3.29	11.80
May	4.32	4.52	-4.63	3.85	10.88
June	4.19	4.32	-3.10	3.81	9.07
Average	4.24	4.38	-3.32	3.94	7.05

Table 1: Monthly values of the average hourly absolute errors of the full range MLP (FR MLP, column 2), the combination of two specific wind range MLPs (SWR MLP, column 3) and the combination of two specific production range MLPs (SPR MLP, column 4). Column 3 gives the improvement's % of the specific wind range models and column 5 gives the improvement's % of the specific production range models.

reasonable for the global and low speed/production models, it might be somewhat large for the high wind/production models, which may result in an overfitting risk.

Table 1 gives monthly values of the average hourly absolute errors of the full range MLP (FR MLP, column 2), the combination of two specific wind range MLPs (SWR MLP, column 3) and the combination of two specific production range MLPs (SPR MLP, column 4). Column 3 gives the improvement's % of the specific wind range models and column 5 gives the improvement's % of the specific production range models. As it can be seen, the specific wind models do not improve on the performance of the full range MLP. Moreover, no high wind models have been built in the months of July and December and the average error of the full range MLP model is 3.32% lower than that of the combination of the wind specific models. On the other hand, the production–specific model combination does indeed improve on the performance of the full range MLP by about 7.05% on the average and, in fact, its performance is better in all months with a gain of al least 10% in four of them.

5 Discussion and Conclusion

The use of specific models for high and low wind speed/production regimes is in principle a sensible idea. For instance, high wind/production situations are much less frequent than their low counterparts and, therefore, a full range model may tend to give more weight to the most frequent regimes. On the other hand, the question of how to define and identify a concrete regime is in itself a difficult one. A simple approach is to do so in terms of a certain wind/production threshold being surpassed. Of course, estimates of the future regime must be based on forecasts. A first idea is to use for that matter NWP wind speed predictions that, in any case, must also be used to derive production forecasts. This may be reasonable for a single farm, but more difficult when, as in the situation described in this work, the production of a large number of farms on a wide area is to be considered. The alternative is to determine the most likely future regime by the production predictions of a concrete full range model; in principle, the problem of deciding whether future productions will be higher than a given threshold should be somewhat easier than that of giving concrete numeric predictions.

In this work we have applied these two approaches to the prediction of the aggregated wind energy production of a very large area of Spain. As our results show, the second, production–threshold based approach, gives consistently better results than the alternative, wind speed based procedure. In any case, more work is needed for a better understanding of the situations underlying production regimes. For instance, a very simple threshold selection procedure has been applied here that, in some cases (at least for wind speed–based regimes) may result in high regimes without actual data; moreover, our high speed/production neural models may have too many parameters and be affected by overfitting problems. Clearly, more thorough procedures should give better results. Another question of interest is the application of these techniques to individual farms. These and other related topics are currently under consideration.

6 Acknowledgments

This work has been partially funded by Spain's TIN 2007–66862 project and the Cátedra IIC en Modelado y Predicción. The second author has been also supported by the FPU-MEC grant AP2006-02285. We thank Red Eléctrica de España for providing aggregated production data.

References

[1] Duda, R., Hart, P., Stork, D. *Pattern classification*. Wiley, 2000.

[2] Hansen, N. and Ostermeier, A., Completely Derandomized Self-Adaptation in Evolution Strategies, Evolutionary Computation 9 (2001), 159–195.